DECISIONMAKING IN COMPLEX MILITARY ENVIRONMENTS

TASK 4

FINAL CONTRACT SUMMARY REPORT

Prepared by:

Gary Klein

Klein Associates Inc.

582 E. Dayton-Yellow Springs Road

Fairborn, OH 45324-3987

Prepared for:

Naval Command, Control and Ocean Surveillance Center

271 Catalina Boulevard

San Diego, CA 92152-5000

Date Submitted: 31 December 1992

ACRONYM LIST

AAW Anti-Air Warfare

AAWC Anti-Air Warfare Coordinator

AIC Air Intercept Coordinator

AU Addition of Utilities

AUD Addition of Utility Differences

CAP Combat Air Patrol

CDM Critical Decision method

CIC Combat Information Center

CIWS Close-in Weapon System

CO Commanding Officer

CoA Course of Action

CON Conjunction

CSE Cognitive Systems Engineering

DCD Decision-Centered Design

DIS Disjunction

DOM Dominance

DSS Decision Support System

EBA Elimination by Aspects

EV Expected Value

HCI Human-Computer Interfaces

IAD International Air Distress

IFF Identify Friend or Foe

LEX Lexiographic

LIC Low Intensity Conflict

LSD Large Screen Display

MAD Military Air Distress

NDM Naturalistic Decision Making

NRaD Navy Research and Development

NSF Number of Superior Features

RPD Recognition-Primed Decision

SA Situation Assessment

SAT Satisficing

SAT+ Satisficing-plus

SEU Subjective Expected Utility

SFD Single Feature Difference

SFI Single Feature Inferiority

SFS Single Feature Superiority

TADMUS Tactical Decision Making Under Stress

TAO Tactical Action Officer

CONTENTS:

EXECUTIVE SUMMARY

The goal of this effort was to examine how the Naturalistic Decision Making (NDM) approach could be applied to the task of designing Human-Computer Interfaces (HCIs) and Decision Support Systems (DSSs). The project was divided into three tasks, each described in a separate report.

Task 1 (Kaempf, Wolf, Thordsen, & Klein, 1992) was to perform a Cognitive Task Analysis of Anti-Air Warfare in the Combat Information Center (CIC) of an AEGIS cruiser engaged in Low Intensity Conflict (LIC) incidents. Fourteen incidents were studied, ten from actual encounters and four from training exercises. Interviews were conducted with decisionmakers, primarily Tactical Action Officers (TAOs), Anti-Air Warfare Coordinators (AAWCs), and Commanding Officers (COs). Both diagnostic and course of action (CoA) decisions were probed. The general finding was that diagnostic decisions were the most critical during these incidents, and that the predominant strategies were feature matching to identify hypotheses about the situation, and mental simulation to construct a causal explanation for complex events such as those where intent must be inferred.

Task 2 (Zsambok, Beach, & Klein, 1992) was a literature review of diagnostic and CoA strategies that might be appropriate for a setting such as the Combat Information Center of an AEGIS cruiser. From these eight diagnostic strategies, three are highlighted in this report: feature matching, analogical reasoning, and mental simulation. This report also highlights a small set of the CoA strategies discussed in Task 2: conjunction, disjunction, single feature inferiority, satisficing, and elimination by aspects. Boundary conditions for the use of these CoA strategies are described. During the project, Tasks 1 and 2 were performed interactively, so that strategies found in the literature shaped the interview probes, and findings from the interviews shaped the literature search.

Task 3 (Miller, Wolf, Thordsen, & Klein, 1992) was the preparation of storyboards for interfaces that should help the CO, TAO, and AAWC handle LIC incidents. The incidents collected during Task 1 were used to synthesize a set of decision requirements that resulted in recommendations about interface features.

The performance of these three tasks suggests the value of a decision-centered design approach. Recent work on Cognitive Systems Engineering (Andriole & Adelman, 1989; Rasmussen, Pejtersen, & Goodstein, 1991; Rouse, 1989; Woods & Roth, 1988) has demonstrated the importance of taking cognitive variables into account for system design, e.g., cognitive workload, attention, and memory. We have shown that decision requirements can be added to this list. It is feasible to identify decision requirements, and to anticipate the decision strategies used by operators performing a task. Furthermore, these analyses can be used to recommend features and configurations of interfaces and decision support systems.

The next step in the project will be to implement the interface features generated through decision requirements, and to evaluate the enhanced displays within an AEGIS simulation

created for the TADMUS project.

DECISIONMAKING IN COMPLEX MILITARY ENVIRONMENTS

OVERVIEW

Introduction

This report is intended to present an overview of the work that Klein Associates has been performing on a project sponsored by the Office of Naval Technology, called Tactical Decision Making Under Stress (TADMUS). TADMUS was designed to learn how Naval officers handle very difficult decisions under conditions such as time pressure and uncertainty. Prior to TADMUS, sophisticated systems were designed to handle high intensity combat conditions. TADMUS was directed at low intensity conflict, including high degrees of ambiguity about the nature of a threat, and the intent of a track. The use of AEGIS cruisers in the Persian Gulf during the Iran-Iraq war was an example of this. AEGIS cruisers were designed for blue-water operations, yet in the Persian Gulf they were needed to operate within very narrow confines, and they lacked some important features for self-defense.

TADMUS is a project aimed at understanding how officers make decisions in a LIC environment, in order to help either with better training of teams and individuals, or with the design of better HCIs or DSSs. Klein Associates began work in support of TADMUS in September, 1990. Our intent was to find ways of designing HCIs and DSSs to improve decisionmaking, building on our past work in Naturalistic Decision Making (NDM) (e.g., Klein, 1989). Previous research on classical, generally analytical, decision strategies has not yielded useful insights for developing better systems for this environment. The question driving this effort was whether a naturalistic decision perspective would do any better.

Our work consisted of three tasks. In Task 1, we conducted interviews with AEGIS commanders and Anti-Air Warfare officers, to study the way they make decisions. The results are described in a separate report (Task 1 Technical Report: Kaempf, Wolf, Thordsen, & Klein, 1992). In Task 2, we surveyed the field of classical and naturalistic decision strategies, to see if there are useful ideas to be incorporated into TADMUS. (Task 2 technical report: Zsambok, Beach, & Klein, 1992). The third task was to draw on both of these efforts to generate a decision-centered approach to designing interfaces and system supports. This task, and the storyboards we developed, are described in a separate report (Task 3 Technical Report: Miller, Wolf, Thordsen, & Klein, 1992). Task 4, reported here, is an overview report of the work conducted in the first three tasks.

The objective of this work, and of the entire TADMUS project, is to generate principles to help decisionmakers overcome acute stressors, in a variety of Navy and other military tasks. To provide context, a specific task was used, that of the AEGIS Combat Information Center (CIC), and the Anti-Air Warfare (AAW) function. This context is described in greater detail in the report by Kaempf et al. (1992). Briefly, the AEGIS cruiser is designed to detect and engage targets at great ranges, particularly threatening aircraft, and to be able to serve as a battle group defense for large numbers of enemy aircraft. AEGIS is designed for large scale, high intensity operations, using its radar capabilities far out at sea to counter enemy raids. The LIC problem is that the AEGIS cruiser has also been called upon for limited objective operations close in to shore. Its system architecture is not well suited to such a mission, creating different types of decision requirements than were originally considered when the cruisers were first developed.

This report is intended to provide an overview of Klein Associates' efforts in the TADMUS program. The report summarizes the major findings of the three more detailed reports, prepared for the three major tasks. The intended audience for this report are the technical professionals involved in the TADMUS project, including the Naval personnel monitoring the project at NRaD, the members of the Technical Advisory Board, and the other researchers working on different aspects of TADMUS. The report may also be of interest to behavioral scientists and professionals working in decision research or studying support system design principles.

Method

The three tasks each employed different methods.

Task 1 was a data-gathering effort to identify the primary types of decisions required for handling the AAW task in an AEGIS CIC. Background interviews were conducted with a variety of personnel in the CIC who were involved in the AAW task, and more in-depth interviews were conducted with COs, TAOs, and AAWCs. These interviews centered around critical incidents and challenging, non-routine events that had actually occurred under LIC conditions. For each event, cognitive probes were used to investigate the reasoning strategies used by the decisionmakers. In addition, some ship-board observations were made to supplement the picture that emerged from the interviews.

Task 2 was a review of the decision research literature, primarily the research on both analytical and nonanalytical strategies. By "analytical," we mean strategies that rely on comparing the pros and cons of options (which is accomplished by evaluating their features). "Nonanalytical" strategies are those described in the naturalistic decisionmaking literature which do not rely on weighing pros and cons of multiple options. Further, they include decisions made not just about options, but also about situation diagnosis.

We evaluated the current state of knowledge of decision strategies, compiling a fairly exhaustive list of candidate strategies and then defining the boundary conditions for each. The intent was to learn when each strategy might be employed, so that designers would know what to anticipate. We also filtered our review through the CIC environment paying particular attention to variables, such as time pressure and information quality, that might rule out the use of certain strategies under operational conditions.

Task 3 was an attempt to synthesize the results of the first two tasks in order to generate storyboards of interface features and decision support concepts that assist CIC personnel in handling difficult and stressful decisions. The results of Task 2 were used to identify decision strategies likely to be used and the results of Task 1 were used to identify the key types of decisions and strategies actually found during operations. These inputs were combined to derive decision requirements for interfaces, and these requirements were used to generate interface concepts. Finally, the interface concepts were represented in a set of storyboards that showed the type of interface features that would follow from a decision- centered approach.

Findings

At this point in the project, after 24 months of effort, we have learned a great deal about NDM in the CIC. The discussion of our findings is organized into eight topics: decision strategies, selecting a course of action, forming a diagnosis, Recognition-Primed decisions, stress and decisionmaking, methods of representation, CIC decisions, and the relationship between Cognitive Systems Engineering and NDM.

1. Decision Strategies

We can distinguish between two types of decisions that would be found in a CIC: Diagnostic decisions about the nature of the situation, and CoA decisions about committing to a reaction to a problem, requirement, or opportunity.

CoA decisions are relatively unimportant in the CIC for AAW. Even though the majority of decision research has addressed CoA decisions, we found that CoA decisions are infrequent, and when they do arise it is for issues that are relatively unimportant. The CoA decisions usually occur when people are working in unfamiliar domains in the CIC, expertise is high and the ability to read the situation is strong. Once the situation has been understood, the reactions are usually obvious.

This finding may seem counter-intuitive, especially since in Task 1 we have studied critical incidents. Many of the events we examined appeared to cover difficult choices, such as whether to shoot down an unidentified aircraft, whether to shoot down Iranian F-4s that appeared menacing, and so forth. However, closer inspection of these decisions showed that the key personnel were not debating the advantages and disadvantages of different courses of action.

Diagnosis decisions are more important than CoA decisions. The primary decision was diagnostic--what was going on. By obtaining a clearer situation assessment, the CoA would become obvious. In the absence of a clear situation assessment, the officers would fashion a CoA that satisfied the need for self-defense while keeping options open for continued diagnosis. For example, there was a maritime patrol incident in which an unknown track ignored radio warnings and kept approaching an AEGIS cruiser. (The track was an unarmed maritime patrol aircraft off the coast of Libya.) The shoot/don't shoot decision was trivial--if the track got within a certain range, it would be shot down. The decision was to judge what type of aircraft it was. The track's identity was sufficiently ambiguous to justify calling in a Combat Air Patrol (CAP) to perform a visual identification, and the tough decision was to judge whether the CAP would arrive in time to perform the identification without screening the aircraft if it needed to be engaged.

2. Selecting a Course of Action

There are a variety of strategies for making CoA decisions. Table 1 lists the 15 analytical CoA strategies we identified in our Task 2 report (Zsambok, Beach, & Klein, 1992). These strategies were taken from the classical decision research literature. In identifying these strategies, we included a number that have been studied in laboratory settings but are not necessarily used in naturalistic settings. They have been posited as logical possibilities (Svenson, 1979).

The 15 strategies in Table 1 are as follows: Expected Value (EV), Subjective Expected Utility (SEU), Addition of Utilities (AU), Addition of Utility Differences (AUD), Dominance (DOM), Conjunction (CON), Disjunction (DIS), Lexicographic (LEX), Elimination by Aspects (EBA), Number of Superior Features (NSF), Single Feature Inferiority (SFI), Single Feature Superiority (SFS), Single Feature Difference (SFD), Satisficing (SAT), and Satisficing Plus (SAT+). These strategies are defined and described in the Task 2 report.

The matrix presented in Table 1 is a summary of the decision strategies most often discussed in the literature. We reviewed these strategies to see which might be relevant for AAW operations in a CIC.

We identified boundary conditions for the CoA strategies. These conditions are also presented in Table 1, related to the type of goal, the information requirement, the method of application, and the environment. Time pressure screens out several strategies that are impossible to perform without time-consuming analyses. Other strategies are intended to find the best option, and are unnecessary in dynamic settings where a satisficing criterion is used. Of the 15 strategies listed in Table 1, five appear to be relevant for the CIC, because they require little time or mental effort. These five are listed in Table 2. Note that even though these strategies have potential for us ein the CIC, the rarity of CoA decisions there makes it unlikely that they would be found during actual incidents.

CoA decisions are hard to identify. There were a few cases where CIC personnel did wrestle with a CoA decision, and we found it difficult to code these according to the decision strategies listed in Table 2. The reasons for the difficulty are instructive.

Table 1 - Boundary Conditions for the use of Decision Strategies for Selecting Among Options

	EV	SEU	AU	AUD	DOM	CON	DIS	LEX	EBA	NSF	SFI	SFS	SFD	SAT	SAT+
GOAL															
Screen									X		X				X

Choose Acceptable						X	X							X	
Choose Best	X	X	X	X	X			X		X		X	X		
INFORMATION:															
Nominal						X	X							X	X
Ordinal+	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Desired	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Undesired	X	X	X	X	X								X		
Uncertain	X	X													
Complete	X	X	X	X	X								X		
APPLICATION:															
Once	X	X	X	X	X	X	X			X	X	X	X	X	

Iterative								X	X						X
2nd Strategy					X	X	X			X	X	X	X		
Relative Eval. -Within	X	X	X												
-Between				X	X			X		X	X	X	X		
Absolute Eval.						X	X		X					X	X
Compensatory	X	X	X	X											
Noncompensatory					X	X	X	X	X	X	X	X	X	X	
ENVIRONMENT:															
Decision Aid	X	X	X	X											
Structured	X	X	X	X	X				X				X		
Time: Little			X		X	X	X	X	X	X	X	X	X	X	

Time: Moderate			X	X					X						X
Time: Extensive	X	X													X

Table 2. Relevant CIC Decision Strategies from the Analysis Literature

(CON) Conjunction

(DIS) Disjunction

(SFI) Single Feature Inferiority

(SAT) Satisficing

(EBA) Elimination by Aspects

One problem is that what counts as a feature is arbitrary. Consider a strategy such as Single Feature Inferiority (SFI), (if an option scores lowest for a feature of interest, it will be rejected). That sounds simple enough. But what is a feature? In the maritime patrol example mentioned above, one option was to use the CAP to perform a visual identification of the unknown track. How would we analyze this option? Do we consider the separate features to be the speed and position of the CAP and the track, or do we combine these into one global feature, that of the CAP getting to the track in time? We don't know if we are dealing with a single feature or not, and the essence of the strategy disappears under our scrutiny. This was one difficulty we encountered in trying to code for the strategies in

Table 1 and Table 2.

A second and related difficulty was that the features became confounded with the decisions. In the maritime patrol example, one feature of the option to use the CAP for visual identification was to judge whether the CAP would arrive in time. This was a difficult judgment, and once it was made, the decision was obvious--if the CAP wouldn't make it, reject the option. If it would, accept the option. If it would be close, use problem solving to work out contingencies. The point is that there was little to be served by judging the CAP's arrival as an input into a CoA decision. It proved difficult to identify "primitive" features, since the context of the decision determined what the features were.

A third difficulty is that when CoA decisions were being made in the CIC, the options were being continually adjusted and adapted as the situation changed. This type of modification is not part of the description of the strategies.

A fourth problem is that in the CIC there is no standard set of evaluation features or dimensions. For one option, certain issues would be relevant, but these would change for a second option. The strategies in Table 2 presuppose a constant set of evaluation dimensions, whereas we found that the dimensions were context-specific and option-specific.

In one incident, an AAWC was trying to find a plane capable of destroying a missile site. He dismissed one option, a strike mission, because it would take too long. He eliminated a second option, using CAP, because of fuel status. The option eventually used, an armed surveillance aircraft, was primarily selected because "It would do the job and was in the area." These shifting evaluation criteria made it difficult to apply the strategies in Table 2.

A fifth difficulty was that the strategies seemed designed for decisions that don't matter very much. They are intended to help a decisionmaker select among options that have very similar utilities, e.g., selecting one house versus another when the features are very close, or the greater strengths of one option are balanced by greater weaknesses. However, in such cases where there isn't a clear favorite, the advantage of choosing the marginally strongest option may be lost in the noise of making the estimates and anticipating how the situation will evolve in the future. The implications of making the best choice may seem important, but the capability of doing so, given the high level of uncertainty, makes the exercise somewhat trivial. This seems particularly relevant to AEGIS CIC decisions where there are high levels of uncertainty and ambiguity, and situations that can change rapidly.

We have identified a CoA decision strategy that does not appear in the literature--the successive pairs strategy. Insofar as we can tell, this has not been described and yet it appears to be a CoA strategy that is relevant for naturalistic settings. The strategy is to conduct successive pair-wise comparisons, choosing the superior of the two options at each comparison point. For instance, if the task is to select one candidate from a pool of candidates, the first two people are compared; the one judged superior is retained and compared to candidate #3, and so forth, until all candidates have been evaluated. The advantage of the strategy is to minimize memory load while thoroughly working through all the options. Only two alternatives are ever represented at any point in time. Every candidate is considered, but for efficiency, the decisionmaker does not compare every candidate to each other. We did not observe the use of the successive pairs strategy in the CIC, and wish merely to identify it as another CoA strategy that might be found in some operational settings.

There is a nonanalytical strategy for option evaluation that probably occurs in naturalistic settings, but receives little attention by researchers--a holistic evaluation. This involves an appraisal of how promising an option appears within the total context, rather than dimension-by-dimension. It is an evaluation that precedes feature examination. For example, a chess player might consider a line of moves and arrive at a possible end-state, which s/he would evaluate in terms of overall potential. Experienced decisionmakers can imagine how a CoA will play out in a situation, and note whether they feel satisfied or uneasy with the position that is emerging. If pushed, the player could easily describe positive and negative features, but many of these features would come after the request to justify the choice, not before it. De Groot (1965/78) has identified this type of nonanalytical evaluation as being important in chess, and it is likely to be relevant to a wide variety of tasks. The idea of a nonanalytical evaluation is difficult to reconcile with an information-processing tradition, but it may be more suited to a connectionist approach to option evaluation.

3. Forming a Diagnosis

Situation assessment and diagnostic decisions are very important in the CIC. The most important decisions were judgments about the nature of the situation, not selections between CoAs. Diagnosis requires that a decisionmaker generate one or more potential hypotheses and evaluate these. Of particular importance were diagnostic decisions to identify a track, to explain observed events, and to determine the intent of that track.

In our review of the decision research literature, we found relatively little on diagnostic decisions and judgments, compared to CoA decisions. We identified eight diagnostic processes or strategies (Zsambok et al., Task 2 report, Table 2) that might occur in the CIC environments. These strategies are listed here in Table 3.

We did not code for all eight strategies, because the identification of these strategies had not been completed at the time the coding was performed. Rather, we analyzed each shift in the decisionmaker's situation assessment and found that virtually all of the assessments involved either feature matching or story generation. Only 1-2% (depending on the coder) fell into the "other" category.

Moreover, some of these eight strategies would not have been possible to identify, anyhow. For example, distinguishing between step-by-step belief updating versus global belief updating would not have been possible because the model contrasts sequential diagnostic judgments with judgments made after all information is received. There was no way to distinguish between these two in the CIC, particularly for retrospective accounts. Others, like mental simulation and story building, are conceptually similar. In this report, we refer to story building as a sub-type of mental simulation. However, there is a substantial literature behind both terms and both are discussed separately in the Task 2 report. Here, we have chosen to define mental simulation as the process of

constructing a causal chain of events that plausibly accounts for an observed phenomenon. It can include the special case of story building, which we define as a type of mental simulation centered around an actor who is carrying out an intent.

Of the eight strategies, two were found in the CIC AAW environment: feature matching and mental simulation (Kaempf et al., Task 1 report). A third strategy, analogical reasoning, has been observed in other operational settings (Klein, 1989; Klein & Calderwood, 1989).

Table 3. Diagnostic Decision Strategies and Processes of Interest to the TADMUS Project

1. Feature Matching

2. Holistic Matching

3. Seeking More Information

4. Story Building

5. Step-by-Step Belief Updating

6. Global Belief Updating

7. Mental Simulation

8. Analolgical Reasoning

We attempted to identify the boundary conditions driving the use of these three strategies. The importance of understanding boundary conditions is to eventually enable designers to anticipate the type of diagnostic strategies people will use in a situation, and to make informed choices about training and supporting these strateggies. In the following paragraphs we present a comparative analysis of these three diagnostic strategies and potential boundary conditions mitigating their use. We also added a fourth strategy for purposes of comparison--Bayesian statistics--which is an analytical procedure for judging hypotheses. Definitions of each of the four strategies are as follows:

Naturalistic Strategies

Feature matching refers to the use of pre-defined and context-free features to adopt an hypothesis or to select between hypotheses.

Mental simulation involves the construction of a story or causal account to infer how a current situation might have evolved from an earlier state (Klein & Crandall, in press). The mental simulation involves construction of a causal chain between the inferred, prior state and the current, observed state. This type of strategy can be used to adopt an hypothesis or to compare different hypotheses. It is suited to domains where there is a high degree of interaction between features, making it difficult to apply straightforward feature-matching methods.

Analogical reasoning involves retrieval of a match to a prior case that can serve to identify the dynamics of the situation.

Analytical Strategies

-

There are a number of possible examples of analytical strategies. The most representative one is Bayesian statistics, in which an hypothesis is evaluated in terms of base rates for different hypotheses and probabilities of the accuracy of different observations. These types of strategies are not available to unaided decisionmakers (e.g., Lopes, 1991), and their value in operational settings has been disputed (e.g., Cohen, in press).

Table 4 presents a matrix showing the boundary conditions for these strategies, parallel to the boundary conditions for CoA decision strategies presented in Table 1. This matrix does not appear in the Task 2 report by Zsambok et al. (1992) which reviews previous literature, rather than presenting new hypotheses. Five dimensions, time pressure, data quality, domain structure, decision timing, and experience, seem to be involved in defining the boundary conditions for the four strategies presented in Table 4.

Time pressure refers to the amount of time available for making the diagnosis. Highly analytical strategies, such as Bayesian statistics, require a great deal of time. Straightforward strategies, such as feature matching and analogical retrieval and matching, require little time. Mental simulation requires conscious effort and is affected to some extent by time pressure.

Data quality refers to the ambiguity and completeness and accuracy of the information available. Again, Bayesian statistics require the highest data quality; if base rates are unknown or unreliable there is little reason to perform a Bayesian analysis. In contrast, analogical inferences can be very creative with inadequate data.

Domain structure refers to the ease with which elements and features can be specified. If it becomes impossible to obtain such a specification, feature matching and Bayesian analyses will be prevented. Furthermore, if there are complex interrelations between elements or important causal linkages, this will further discourage feature matching and Bayesian statistics.

The timing of decisions refers to whether diagnoses are required as each new datum is received, or after all the evidence has come in. This can affect the use of belief-updating strategies.

Finally, the experience level of the decisionmaker (e.g., familiarity with the domain) will affect the choice of strategy. Novices are more comfortable using feature matching, especially if the features are defined for them, although they may have difficulty judging when a feature is present. Similarly, novices can easily use analogues, although they may

Table 4. Matrix of Diagnostic Strategies and Boundary Conditions

	Diagnosis Strategy			
	Feature Matching	Mental Simulation	Analogical Reasoning	Bayesian
Time needed				
Low	x		x	

Medium		x		
High				x
Data Quality				
Low			x	
Medium	x	x		
High				x
Domain Structure				
Low		x	x	
Medium				
High	x			x

Decision Timing				
Updating	x		x	x
Terminal		x		
Experience Level Needed				
Low	x		x	
Moderate		x		x
High				

draw the wrong conclusions. More experience is needed to construct an adequate mental simulation, as well as to select and use base rates in a Bayesian analysis.

Strategies used in the CIC

As reported in the Task 1 report (Kaempf et al., 1992), the most frequent diagnostic strategy was feature matching, but the use of mental simulation emerged as a critical strategy for many of the difficult decisions. They found that 88% of the diagnostic decisions were coded as showing feature matching, and 11% were coded for mental simulation. As discussed above, one way to think of mental simulation for diagnosis decisions is the

construction of a story to explain how a state of events might have been caused. We can expand on these findings as follows.

Feature matching was the dominant diagnostic strategy. In most settings, there are a clear and limited set of features to consider. The track's target is unknown, but it has taken off from a hostile country, it is flying outside the pre-defined commercial air corridors, it is emitting Identify Friend or Foe (IFF) signals associated with fighter aircraft, it is heading directly at a friendly ship, it is ignoring radio warnings to alter course, it is using its fire control radar to lock onto the ship, it is flying at a very low altitude, which enables it to avoid detection and attack by a ship, and it is flying very fast. Perhaps it is also showing a hostile flight pattern--flying low, then popping up for a brief lock-on, then returning to a low altitude. This set of features is sufficient to worry most Naval commanders.

One interviewee told us that of these five features (country of origin, IFF, bearing, response to radio warnings, end attitude), if three of the five are threat-related the track is treated as a threat. While this may sound straightforward, further questioning showed that a single feature, Mode II IFF, would be considered a sufficient indicator of a hostile aircraft.

While the use of mental simulation appeared in only 11% of the diagnostic decisions, this figure may be misleading, and we feel that the actual rate was higher. We have adopted very strict criteria for coding the use of mental simulation. The ambiguity of the incident accounts makes it very easy to infer the use of mental simulation, and we wanted to be sure we did not manufacture evidence for it. There were many cases where coders believed that the decisionmakers had to be constructing stories to account for events, or to generate expectancies. They did not, however, classify these as mental simulation unless the incident account specifically included a reference to inferring a causal sequence, imagining how a course of events might have unfolded, or placing oneself in the position of an adversary to infer what the adversary might have been planning. Absent this type of reference, we used the more conservative category of feature matching to denote simply that the decisionmaker used a set of features to make a diagnosis.

Mental simulation shaped the diagnoses for some of the most difficult decisions. These were judgments about the identity of a track, and the intent of the pilot flying the aircraft. In these types of instances, there were data to be integrated, rather than just features to be checked off. The mental simulation, usually in the form of building a story, provided a causal link between different observations. For example, in an incident in which Iranian F-4s had taken off and were circling near an AEGIS cruiser, the CO of the cruiser used their flight paths, radar activities, and so forth to build a plausible story of how they were just harassing him. The different observations fit this story well. Another possible story was that they were preparing to attack him, which was also plausible since they had turned on their fire control radar to lock on to his ship. However, the CO did not believe this second story, since the behavior of the F-4s was so brazen, so attention-gathering, that he could not imagine a serious pilot preparing an attack in this way. The CO needed to prepare for an attack, and did so, but he held his fire, despite provocation, since he did not believe that the attack story was plausible.

4. Recognition-Primed Decisions

One of our goals in this project was to study how the Recognition-Primed Decision (RPD) model (Klein, 1989) mapped onto the activities in the AEGIS CIC. We found that the RPD model provided a good account, since it emphasizes the recognition of situational dynamics as one of the key drivers of the selection of a CoA, and this was what we observed. The RPD model describes how decisionmakers can adopt reasonable CoAs without having to compare alternatives, and this also fit our observations.

We were able to elaborate on the RPD model, using the data we collected in our interviews. Figure 1 shows the complex version of the RPD model. In non-routine incidents, the decisionmaker may have difficulty sizing up a situation, and may need to obtain more information or may need to think more about the cues. Once the situation is understood, the decisionmaker recognizes plausible goals, as well as feasible CoAs, and can select one CoA by this process: appraising the CoAs in rough order of their potential, choosing the first one that appears to work, and then evaluating it using mental simulation to

consider how it will play out in context.

As a result of this research project, we have also deepened our account of what happens in Figure 1. It may be worth describing how our account has expanded, since the RPD model is the basis of the storyboards we developed for an improved interface.

First, the decisionmaker sizes up the situation, categorizing it as familiar based on experience. Previously, we had nothing more to say about this process. Now, we would add that the recognition of familiarity can take several routes. (a) There can be a recognition based on feature matching. (b) The person can recognize the pattern of events as fitting a familiar story. (c) The person may use a link to an analogous event to recognize the situation. (d) The person may use a holistic recognition strategy, e.g., holographic memory retrieval or a connectionist account that does not rely on features.

If the decisionmaker is unable to recognize familiarity, then one reaction is to seek more information, which in turn changes the experience, as shown in Figure 1.

Figure 1. Complex RPD Strategy

The decisionmaker can also consciously attend to the diagnosis, either by deliberately noting the features, or by constructing a story linking the features. If there are several possible stories or causal explanations, the decisionmaker may consciously appraise each for

plausibility, consistency, and so on. This process has been described as the story model of Pennington and Hastie (in press) and the process of mental simulation described by Klein and Crandall (in press). Based on the story building/mental simulation, the decisionmaker may still request more information, or may arrive at a serviceable diagnosis.

Mental simulation also enters the RPD model once the diagnosis is made, since the simulation will also enable the decisionmaker to form expectancies. The mental simulation provides a model of how the situation should unfold, if the simulation is accurate (and the

explanation of events is correct). This process generates expectancies. If the expectancies are fulfilled, confidence in the mental simulation goes up. If there are surprises, the decisionmaker may try to explain away the disconfirming evidence, thereby adding more complexity to the mental simulation (see Cohen, et al 1992; Klein & Crandall, in press for a discussion of this process). Further, the expectancies help decisionmakers prepare to react to events.

The last elaboration of the RPD model involves the evaluation of CoAs. Mental simulation for diagnosis postulated a previous state of affairs that was projected into the present to try to account for the current situation. In evaluating a CoA, mental simulation is used to project into the future. While we retain this use of mental simulation for option evaluation, we also encountered a number of cases in Task 1 (Kaempf et al., 1992) in which the evaluation was much more straightforward. The decisionmaker recognized a reasonable CoA and assessed whether it satisfied a small set of criteria. This is a form of feature matching, without invoking causal linkages. In fact, this is a clear case of satisficing, and so we include it as a second type of CoA evaluation that does not require comparison between CoAs. A third type of CoA evaluation might involve analogical reasoning, whereby a

decisionmaker assesses the feasibility of a CoA by retrieving an analogue where a similar CoA had been successful.

The results of Task 1 showed that the CIC decisionmaking for AAW predominantly fit the RPD model. We did not code using "RPD" as a category since we wanted to study the components of the model--feature matching and mental simulation. These were the dominant categories observed. There were very few cases where decisionmakers explicitly compared one option to another, which is in line with recognitional decisionmaking.

5. Stress and Decisionmaking

Our literature review in Task 2 did not specifically examine the range of stressors that might be found in a CIC. Nevertheless, we have recently performed an additional literature review to determine how acute stressors might affect decisionmaking (Klein & Zsambok, in preparation).

We identified several acute stressors: crowding, noise, performance pressure, workload, anticipated pain, anticipated danger, and emergency conditions. Our conclusions were based primarily on a number of recent summaries of the research literature on stress.

We found that acute stressors appear to mediate cognitive performance in a few ways:

· by changing speed/accuracy tradeoffs

· by creating a distracting secondary task of managing the stressor

· by taking up space in working memory

· by restricting the set of cues that can be physically attended to

· by increasing fixation on the first option considered since confusion can arise when

several options must be simultaneously considered

· by increasing conservatism by making the decisionmaker reluctant to adopt a

complex CoA that will be difficult to enact under time pressure and distractions.

All of these reactions appear reasonable in the face of the acute stressors, and do

not seem to constitute biases.

The effect of acute stressors is to make people less likely to use analytical, compensatory strategies, since these are difficult to carry out even during non-stressed conditions. Instead, naturalistic strategies are likely to be found, such as recognitional decisionmaking, since the decisionmaker wants to use experience to quickly arrive at a reasonable diagnosis or CoA.

6. Methods of Representation

There is much more to studying decisionmaking in the AEGIS CIC than categorizing the types of strategies. It is perhaps more important to be able to account for the line of reasoning that enabled a commander to diagnose a situation or adopt a CoA and to understand how a strategy is carried out.

For purposes of design, it is more important to represent the line of reasoning than to merely label the strategy. It is our assumption that by understanding the thinking processes of key decisionmakers, we will best be able to inform designers on how to configure HCIs and provide DSSs.

This is one of the important premises of Cognitive Systems Engineering--that by understanding cognitive processes we will be able to design better HCIs and DSSs. For our TADMUS project, we sought to find ways to address one type of cognitive process--decisionmaking. This section describes the methods we derived for representing the reasoning and thinking that underlies decisionmaking during AAW operations in an AEGIS CIC. At some point we may want to speculate about how these representations should be configured so they can be used by design engineers.

Cognitive Task Analysis. Much of our work in this project was centered around getting inside the heads of TAOs and COs as they made difficult decisions during non-routine and challenging incidents. We adopted the critical incident framework because our previous research (e.g., Klein, Calderwood, & MacGregor, 1989) showed that expertise emerges most clearly during these events than during routine tasks. Furthermore, by studying actual events we can learn about different types of context that might be hidden by such other methods as providing subjects with simulated tasks since simulations may be unable to incorporate certain important features of situations. The down-side of a critical incident method is that it relies on memory for events that may have taken place months or years previously. We have found that the non-routine incidents are remembered surprisingly well, but we always need to be cautious since the informants' memories may not be reliable, and they may construct

details. We try to catch inconsistencies, but faulty memory is an inevitability that must be acknowledged.

Our approach to Cognitive Task Analysis in Task 1 was to use concept maps (Gowin & Nowak, 1984; McFarren, 1987) as a first attempt to gain a global understanding of the domain and of the key decisions. Then we shifted to using the Critical Decision method (Klein et al., 1989) for more detailed probes into the diagnostic and CoA decisions, the cues and information used, the nature of the reasoning, the tacit knowledge and perceptual distinctions, and so on.

There are other methods for Cognitive Task Analysis that can be used. Think-aloud protocols are perhaps the most popular method, but we have avoided them because they take a subject through an artificial, rather than an actual, incident. If the task is abstract, such as cryptarithmetic, then there is no problem. But for a concrete domain such as AAW operations, then context is very important and artificial problems can be misleading. In most of the incidents we studied, some type of contextual cue generally seemed to enter, whether it was incidental knowledge about the maintenance problems with an adversary's radar, or lack of confidence in the crew of one of the ships, or the inadequacy of briefings during a change in command.

The Critical Decision method was applicable for this domain. Even for incidents that had occurred several years in the past, the participants had a clear memory and were able to provide a good account of what they had been thinking about.

Representational formats. We used several formats in analyzing the data. First, the interviews had been audiotaped. Second, we obtained a transcript for each tape. The transcripts were usually between 40-50 pages in length. Third, we derived brief accounts (two to four pages) of the essentials of the incident. Fourth, we codified the transcripts and interview notes using a format illustrated in Figure 2 for the "Harassing F-4s" incident

described earlier. Fifth, we developed a graphic representation of the incident, as illustrated in Figure 3 for the event. This Decision Flow Diagram stands in contrast to a Data Flow Diagram, which is a stock in trade for software developers, specifying the types of data transactions that the system must accommodate. The problem with a Data Flow Diagram is

that it does not show how the user is thinking about the task, and so is not helpful in designing a useable display. In contrast, the Decision Flow Diagram, and the brief incident account, are intended to help a designer understand the context in which the system will be operated.

ProblemNo.	Problem Description	Type	SA and CoA Description	Strategies
1	Must determine intent of unknown tracks.	SA 1	Typical Iranian patrol aircraft. Of interest because of recent hostilities.	Generate Story Single Story NA
2	Must deal with tracks that are more hostile.	SA 1	Lead aircraft heading toward cruiser. Has fire control radar on. More of a potential threat.	Feature Match Moderately Familiar Uncoded
3	Tracks become increasingly more hostile.	SA 1	Lead aircraft expanding its orbit and getting closer. Intentions appear more hostile. Must let tracks know the cruiser will not tolerate.	Feature Match Highly Familiar NA

		CoA 1	Warn tracks on comm. nets.	Recognize No Evaluation NA
		CoA 2	Break lock of fire control radar.	Recognize No Evaluation NA
		CoA 3	Prepare to illuminate tracks.	Recognize Evaluate Choose via Mental Simulation
4	Tracks have gotten too close.	SA 1	Tracks are probably harrassing. Turning toward the cruiser will indicate hostile intent.	Generate Story Single Story NA
		CoA 1	Prepare to engage if the if the track closes the cruiser.	Recognize Evalualte Choose via Mental Simulation

5	Must determine tracks intent.	SA 1	Tracks are still orbiting. Stop using fire control radar. Tracks appear to be less threatening.	Feature Match Highly Familiar NA

Figure 2. The Harassing F-4s: Incident Account

Figure 3. Harassing F-4s: Decision Flow Diagram

7. CIC Decisions

We obtained some information about the nature of CIC decisions, which may have some general uses for TADMUS and for other projects as well.

Goals and functions of anti-air warfare in the AEGIS CIC. We reviewed the 14 incidents that had been studied, and identified a set of 15 common goals and functions. These are listed in Table 5. They include identifying a track, determining the intent of the

pilot of an aircraft, avoiding escalation, preparing to engage, and so forth. These are the difficult tasks that make up anti-air warfare, at the level of the commander or the air TAO. HCIs and DSSs need to help operators in the CIC perform these tasks.

From Task 1, we were able to form some idea of how these functions are performed. Track identification (F) and intent determination (A) often require story building or mental simulation to link together the different data elements into a diagnosis. Preparing to engage (D and H) requires a clear visualization of factors such as effective firing range of the AEGIS cruiser, and the target. It also requires the crew to be sensitive to timing issues.

Recognition of a problem (B) usually relies on feature matching to determine if the ship is vulnerable to a track that might choose to attack. Trying to avoid escalation (C) involves problem solving and mental simulation to find a defensive CoA and anticipate whether it might provoke an attack.

All of these goal states were found in the incidents we studied. We also could anticipate the decision strategies used by going back to the incident accounts themselves. This type of analysis provided inputs into the HCI design which attempted to support the key decisions and goals.

It is possible that the goals/functions listed in Table 5 will be relevant to U.S. Navy missions beyond the AEGIS CIC Anti-Air Warfare domain. The identification of a track appears to be a common task, and the strategies such as feature matching and mental simulation that we have observed may be common to other domains as well. Determining intent also seems to be a common requirement, and again a strategy such as mental simulation (story building) may generally be used to synthesize different types of data. If so, then the display concepts emerging from this study may have broader applicability.

Table 5. Primary Goal States of the Critical Incidents

A. Determine intent: CIC crew attempts to determine the intentions of a track, such as whether or not the track is hostile.

B. Recognition of a problem: crew tries to determine if they are faced with a potentially threatening situation.

C. Take actions to avoid escalation: crew takes deliberate steps to avoid the escalation of an incident into an engagement.

D. Take actions toward engaging track(s): crew takes preparatory steps needed to engage a track.

E. Monitor on-going situation: the CIC crew monitors a situation to detect any changes in the situation.

F. Identify track: crew attempts to determine the identity (e.g. country of origin) of a track.

G. Allocate resources: the CIC crew attempts to allocate limited resources to deal with the current situation.

H. Prepare self-defense: crew takes steps toward self-defense, such as bringing up the CIWS.

I. Conduct all-out engagement: crew actively engages a track with a weapon system.

J. Monitor tracks of interest: crew monitors a tack which has some significance to the current situation.

K. Reset resources: the crew returns ship resources to pre-incident status.

L. Collect intelligence: CIC crew actively tries to collect information on a track.

M. Trouble-shoot: crew tries to trouble-shoot a system.

N. Determine location: CIC crew attempts to determine the location of a reported track.

O. Other: goals not coded in the above list.

AEGIS recommendations. We have made two observations that may be relevant to AEGIS operations.

First, there is a CIC goal/function to avoid escalation. This is typical of low intensity conflict, and occurs in many similar settings. We have seen four different procedures for avoiding escalation: radio, CAP, breaklock, and illumination. When there is a problem, the first reaction is to try to achieve radio contact, using IAD or MAD. Another reaction is to

call for CAP to perform a visual identification, and, if necessary, to escort the track away. A third reaction is to breaklock, such as occurred in one incident in which Iranian F-4s locked up on an AEGIS cruiser. The commander didn't want to illuminate the F-4s, for fear of surprising and panicking them into doing something rash such as releasing a missile, so he just broke the lock of the F-4s' fire control radar, to let them know they had been observed, and to prevent them from taking hostile action. A fourth reaction is to illuminate the track, in order to warn of the immediate danger of being engaged.

Our recommendation concerns this fourth reaction. It only applies to military aircraft, since commercial airliners don't have the equipment to pick up the signal. But for military aircraft, if the goal is deconfliction, it is important to convey that the airplane is in danger, prior to engagement. We found that some commanders were reluctant to illuminate a track, since the process of illumination compromised their ability to use AEGIS effectively. The odd result is that a powerful message, illumination, which was available to commanders of ships with older technology, was less readily available to AEGIS. We recommend that the AEGIS system be reconfigured to allow commanders to use illumination without having to take over manual control. That is, the illumination request should be made through the computer, and entered into the queue of tasks.

The second recommendation concerns the flow of information in an AEGIS cruiser. The idea appears to be to defer decisionmaking initiative to the lowest levels, so that rapid responses can be made without having to pass decisions up and down through channels. This makes perfect sense during overload conditions. It makes less sense during low intensity conflicts. For many of the incidents we studied there were few tracks of interest, sometimes only one. So there was excess capacity in the CIC, with opportunities to reassign tasks and to request specialized assignments (e.g., crew member "x" is assigned to monitor the track's heading and to inform the captain if it deviates suddenly). The ambiguity of low intensity conflicts also differentiated them from a high intensity engagement. Identity and intent of tracks were harder to estimate, and there was a stronger need to synchronize situation assessment. It often seemed that here the information kept flowing up to the TAO and CO, whose job it was to put the pieces together. But the TAO and CO sometimes appeared to neglect to communicate their situation assessment to the rest of the crew in the CIC. For instance, an Air Intercept Coordinator (AIC) might use CAP differently when the track has positively been identified as hostile than when some suspicious signals have been received. Currently, the AIC is generally left to draw his own conclusion, rather than having the commander let him know the current status of the track's identity and intent. We recommend that alternative means of coordination be considered for low intensity conflicts.

8. Cognitive Systems Engineering and Naturalistic Decision Making (NDM)

The purpose of our TADMUS project was to explore and demonstrate the usefulness of a NDM framework for designing interfaces and decision support systems. This effort falls within the field of Cognitive Systems Engineering, which is the attempt to use our understanding of cognitive processes to support the design process. Cognitive Systems Engineering tries to take into account factors such as memory and attention and workload and situation assessment and decisionmaking (Rouse, 1989; Woods & Roth, 1988). In this project we have tried to develop and apply a specific method for identifying decision requirements and using them to drive the design. If we were successful, we would show that decision requirements could be incorporated as a component of Cognitive Systems Engineering, and the method itself would be one example of how to proceed.

This section of the report reviews the goals and results of Task 3 of our contract, to develop storyboards for the CIC, for AAW.

The goal of Task 3 was to use our knowledge of NDM strategies, particularly strategies such as the RPD model, to generate concepts for a human-computer interface, and for decision support. A second goal was to identify some principles for taking decision requirements into account in deriving interface concepts and storyboards. Both of these goals were successfully achieved.

The Task 3 report (Miller, Wolf, Thordsen, & Klein, 1992) presents the full set of storyboards and details about methodology. For this overview, only a few storyboards will be described, along with the methods we used.

The decision flow diagrams and formatted representational sheets can be used to define the primary decisions,(see Figures 2 and 3). These goals and functions also are the key decision types to be made in this environment, and constitute the primary decision requirements for designers to satisfy. In this way, NDM can be used to make key decisions the focus of the design process.

The primary decisions and functions can become the focus of HCI and DSS requirements. We derived a method for combining all instances of a function, such as determining intent, across the different incidents in which it occurred. We then identified the critical cues and patterns used to make this judgment. These became the necessary inputs, and further helped to shape the specifications that would be given to designers. The process is illustrated in Figure 4. The incidents are arranged in columns, and the decision goals/functions are represented as linked to each incident. For example, the goal of determining intent (A) was relevant for five incidents. To make it easier to see how the crew members determined intent, we could go back to the five incidents to see which specific cues and relationships were used to make the inferences.

Therefore, we evolved a procedure for going from the incident accounts to the specification of necessary cues, including the description of relationships between cues. We use the term "decision requirements" to refer to the goals and decisions themselves, and also to the constellation of cues and decision strategies, all of which must be understood in order to provide useful support.

For example, in preparing to engage, we found that the traditional AEGIS concept of striking the adversary at the greatest possible distance no longer applied. Instead, during low intensity conflicts, the commanders moved into the opposite task--letting the track approach as close as possible, firing at the last possible moment, in the hopes of avoiding an engagement. Further, we found that commanders had learned to rely on subtle cues, such as the fact that an enemy aircraft was likely to turn away after firing a missile, and that this turning away was seen more quickly on the displays than was the missile itself. The turning away pattern therefore has become an early indicator of a missile. So, what is needed is a display that showed, not only the position of the track, not only its heading, but changes in the heading. Changes in heading represent the first derivative. For other features such as velocity, changes in the rate of change are needed. This represents the second derivative, acceleration. A display that shows acceleration can provide the operator with greater sensitivity to the dynamics of the incident.

Another type of cue needed to help CIC crew members prepare to engage (goal state D), was the dynamically changing boundary between vulnerability and lethality. If a commander is to be able to wait as long as possible, it is necessary to visualize the range of the enemy's missiles, along with the AEGIS cruiser's capability. This suggests HCI and

DSS graphics features that would enable a commander to wait for the last possible moment, without endangering a ship. We also heard in a number of incidents about engagements that were avoided with less than a minute to spare, or captains with a greater experience base allowing threats to approach more closely than TAOs would tolerate, thereby avoiding engagements with tracks that turned out to be non-threats. Figure 5 shows a storyboard that portrays the weapons release ranges of the ownship and the target. This can give the commander a sense of where the last possible moment is, in order to preserve options as long as possible.

Figure 6 shows a storyboard for illustrating the ship defenses, such as CIWS. During a Low Intensity Conflict, the commander will not want to bring up these systems too early, but must be sure they are ready to be used at the critical moment. During periods of intense communications it was not always clear to the decisionmakers we studied what was the readiness state of their weapons systems. Figure 6 could help provide this information.

Other storyboard concepts presented windows of opportunity, to alert commanders as to the timing for taking certain actions. A related concept was the use of the interface to set

tripwires, especially during the preparation for the mission. These tripwires could function as context-sensitive alerts to key features of the mission, by serving as a tap on the shoulder.

Returning to the decision requirement of determining intent, we found that this often depended on story building, and the stories were often tailored to the specific events during the incident.

We used our knowledge of naturalistic decision strategies to conceptualize types of HCI features and DSSs. Specifically, the use of mental simulation including stories for

situation diagnosis suggests interfaces that enable the decisionmaker to build a story, and possibly even interfaces that portray the story. For example, historical displays would enable a commander to review the development of the incident. These displays might need to show critical events, such as when attack radars were used, rather than just heading and speed. At times, it might be useful to link up the histories of different tracks, to see if they are part of a coordinated attack.

We suggested an interface that portrayed the history of the incident, as shown in Figure 7. The current AEGIS system has some capability to portray history, but some important details such as when a track changes course, turned on fire control radar, deviated from its course, and so forth are lost. Figure 7 shows how it could be possible to review the target's history in order to derive inferences about intent.

One of the weaknesses of a story building strategy is that it can be heavily dependent on recent details, as initial details that might turn out to be significant are forgotten or were never noticed. Alternately, initial details can color the way the story is built, despite disconfirming evidence. By giving the commander the capability to review the entire incident it may be possible to improve the accuracy of the story building, and to speed up the realization that the original story was wrong.

The use of display concepts to support mental simulation and story building could have application for tracking expectancies, and for evaluating proposed courses of action, and might even be relevant to other domains such as diagnosing causes of malfunctions in nuclear power plants. It is difficult to infer causes from the cues available, since they can become tangled;

the operator really needs a picture of how the malfunction emerged over time, in order to build a story about what is going wrong. Therefore, one outcome of this work is to suggest concepts for diagnostic displays in general.

The design of the AEGIS system was driven by the need to defend during High Intensity Conflicts, with large numbers of friendly and hostile tracks on the screens. There would be little value in providing further clutter using the interface features shown in Figures 5, 6, and 7. However, we found that during Low Intensity Conflict, clutter was not a problem, particularly as there are four large screen displays in a CIC, and one could easily be configured to be dedicated to a track of interest. Therefore, we are not implicitly criticizing the current AEGIS screens and interfaces, since they were not developed to address Low Intensity Conflicts.

The storyboards shown in Figures 5, 6, and 7 show that the assessment of critical incidents could lead to a straightforward identification of decision requirements, linked to the goals listed in Table 5. The incident accounts themselves stimulated design concepts. Our understanding of the decision strategies used to accomplish different goals further informed our thinking about how to represent information. On one level, we noted that most inferences were drawn using feature matching and mental simulation (particularly story building), rather than comparing courses of action. We speculate that for other missions, such as planning, we would see more comparison between courses of action.

At a deeper level, we could connect decision requirements to strategies. This has implications for the way the decision could be supported. However, we had to be careful not to be too simplistic about these categories. We noted that in pursuing a goal, individuals often switched from goal state A to goal state B, to C, back to A, and so forth. Moreover, some goal states operate concurrently. The interface needs to allow for the interaction between goals, primarily by the way it highlights critical cues and factors. As the operator changes goal states from, e.g., determining intent to preparing to engage, the cues that are critical change as well. The interface would hopefully take this into account so the operator is given the relevant information in the midst of a dynamic incident.

Conclusions

The use of NDM as a focus for Cognitive Systems Engineering seems to be promising. At this point in the project we have a much clearer understanding of the naturalistic decision strategies. We were able to use cognitive task analysis to identify key decisions, to determine the decision strategies used, and to provide a picture of the reasoning processes used in carrying out those decision strategies. And we are continuing to develop methods for transforming a knowledge of decisionmaking into specifications for interface features and decision support systems.

Some of the most likely decision strategies to expect in naturalistic settings have been identified and their boundary conditions have been described. We can distinguish between diagnosis decisions and CoA decisions. For diagnosis decisions, we have defined a small set of likely strategies. Furthermore, we understand some of the boundary conditions for these strategies--time pressure, the timing of the diagnoses, and the complexity of interrelationships between cues. In well-structured domains, for decisions involving small and independent sets of features, we expect that feature matching will be sufficient. As the interrelationships increase, and the importance of causal linkages increases, then story building becomes more important, and displays are needed to permit the decisionmakers to formulate these stories and assess their plausibility. For CoA decisions, factors such as data quality and time pressure primarily distinguish between the compensatory and noncompensatory strategies. Many of the noncompensatory single-option strategies could be used in the CIC type of environment, but there seems little payoff to tailoring interface features to them, since few CoA decisions are made, and these do not appear to be critical. Moreover, singular strategies such as mental simulation appear to be effective in enabling crew members to use experience to adopt a CoA, without having to compare options. There is more likelihood of comparing options during mission planning than during operations.

In general, naturalistic decision strategies such as the RPD model appear to do a good job of describing the way commanders actually work under the stress of operations, in non-routine incidents. We have elaborated components of the RPD model, as we have learned more about diagnostic decisionmaking. In particular, the process of mental simulation seems to serve a variety of important functions, for diagnosis, providing expectancies, and CoA evaluation.

Traditional approaches to display design have relied on task analyses and data flow diagrams to identify cues that are necessary on a screen. Unfortunately, these analyses have difficulty in portraying to designers how the cues are used, what interrelationships are important, and how the decisions are being made. Skilled designers will try to ferret this information out themselves, but under time pressure, many designers just want to be shown the specifications, without having to do additional work to take on the perspective of the decisionmaker.

A Cognitive Systems Engineering approach takes the opposite view--that there will be an important payoff if we can represent the cognitive processes of the user to the designer. This is a large challenge, particularly since the field of cognitive psychology is fairly broad. The chore of describing a full array of cognitive processes (including memory, attention, workload, perception, learning, and so forth) seems inefficient. The analyses may take too long, the payoff is unclear, and the methods for capturing and portraying a full array of cognitive processes are still unproven, even in areas that have been well studied such as workload.

The premise of this project was that a more directed study of decisionmaking might offer greater potential than a full CSE approach. If we can learn to use what we know of NDM, we may be able to assist designers with greater efficiency. Instead of a task analysis and a set of data flow diagrams, we performed a cognitive task analysis, and derived a set of decision flow diagrams. The resulting HCI recommendations were different from the displays usually found in CICs. They highlight story building for diagnoses. They focus on specific cue interrelationships, and on sensitive changes in cues, such as second derivative cues to portray dynamic events such as sudden changes in a track's altitude or bearing. Currently, the AEGIS display does not even present first derivatives, i.e., simple trends. It presents altitude and speed as digital data on a separate display, rather than presenting 2nd derivative cues. Our goal is to maximize the sensitivity of the interfaces so that the reaction time of the CIC crew members is shortened, and their ability to control events is increased, allowing them more time in making diagnoses. In fairness, we must note that the designers of the current AEGIS system were attempting to maximize responsiveness to major threats, rather than increase precision during low-intensity events.

There is a risk in using decision strategies to design displays, since the strategies observed may be specific to the last generation equipment, whereas the new displays may themselves transform the decision strategies. We need to be careful in handling this problem, since it is definitely a limiting feature of a NDM approach to Cognitive Systems Engineering. Nevertheless, it seems unlikely that key decision requirements such as determining intent will radically change, or that a story-building approach to synthesizing causal relationships will cease to be relevant. We may find that these super-ordinate decision types and strategies may remain more stable than data flows and task analyses.

It may also be useful to contrast a Cognitive Systems Engineering approach, using NDM, with an ecological approach to design (e.g., Flach, Hancock, Caird, & Vicente, in press). An ecological approach attempts to define and represent invariants in the environment, without recourse to the cognitive processes of the user. This is an interesting and important strategy, and such displays should help users see directly what is going on during an incident. However, there are many invariants in the environment, and the choice of invariants to portray, plus the form of their display, are not easily resolved. It is our hope that the design process will be strengthened by taking the additional step of incorporating the user's cognitive processes, i.e., decisionmaking activities, into account.

A Cognitive Systems Engineering approach centered around naturalistic decisionmaking may be useful for organizing related design concepts. In addition to helping the designer visualize which decisions the system must support, how they are going to be made, using which strategies, the cognitive task analysis also highlights the critical cues and relationships, linking these together as a function of the type of decision they must support.

The next steps of this project are to take these ideas forward, first by evaluating and expanding on the storyboard concepts already developed, and then by implementing these concepts in the architecture of a decision support tool being developed at NCCOSC. In this way we can evaluate the effectiveness of these concepts and this approach to using decisionmaking to guide the design process.

References

Andriole, S. J., & Adelman, A. (1989). Cognitive engineering of advanced information technology for Air Force systems design & development: Prototypes for air defense intelligence & operations. Marshall, VA: International Information Systems, Inc.

Cohen, M. S. (in press). The naturalistic basis of decision biases. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Cohen, M. S., Adelman, L., Tolcott, M., Bresnick, T., & Marvin, F. (1992). Recognition and metacognition in commanders situation understanding. CTI Technical report (1992). Arlington, VA.

de Groot, A. D. (1965/1978). Thought and choice in chess (2nd ed.). NY: Mouton.

Flach, J., Hancock, P., Caird, J., & Vicente, K. (Eds.). (In press). The ecology of human-machine systems. Hillsdale, NJ: Lawrence Erlbaum Associates.

Gowin, D. B., & Novak, J. D. (1984). Learning to learn. NY: Cambridge University Press.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. Bower (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 18, pp 1-47.

Kaempf, G. L., Wolf, S., Thordsen, M. L., & Klein, G. (l992). Decisionmaking in theAEGIS combat information center. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

Klein, G. A. (1989). Recognition-primed decisions. In W.B. Rouse (Ed.), Advances inMan-Machine System Research, 5, 47-92. Greenwich, CT: JAI Press, Inc.

Klein, G. A., & Calderwood, R. (1988). How do people use analogues to make decisions? Proceedings of Case-Based Reasoning Workshop, 209-223.

Klein, G. A., Calderwood, R., & Mac Gregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics, 19(3), 462-472.

Klein, G. A., & Crandall, B. W. (In press). The role of mental simulation in naturalistic decision making. In J. Flach, P. Hancock, J. Caird, and K. Vicente (Eds.), The ecologyof human-machine systems. Hillsdale, NJ: Lawrence Erlbaum Associates.

Klein, G., & Zsambok, C. E. (In preparation). The effect of acute stressors on decision making. In J. E. Driskell and E. Salas (Eds.), Stress and human performance. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lopes, L. L. (1991). The rhetoric of irrationality. Theory & Psychology, 1(1), 65-82.

McFarren, M. R. (1987). Using concept mapping to define problems and identify key kernels during the development of a decision support system. Master's thesis, School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH.

Miller, T., Wolf, S., Thordsen, M. L., & Klein, G. (1992).A decision-centered approach to storyboarding anti-air warfare interfaces. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

Pennington, N., & Hastie, R. (in press). A theory of explanation-based decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. (1991). Cognitive engineering:Concepts and applications. Wiley & Sons.

Rouse, W. B. (Ed.). (1989). Advances in man-machine systems research, 5. Greenwich, CT: JAI Press.

Svenson, O. (1979). Process descriptions in decision making. Organizational Behavior and Human Performance, 23, 86-112.

Woods, D. D., & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. Human Factors, 30, 415-430.

Zsambok, C. E., Beach, L. R., & Klein, G. (1992). Decision-making strategies in theAEGIS combat information center. Fairborn, OH: Klein Associates Inc. Prepared under contract N66001-90-C-6023 for the Naval Command, Control and Ocean Surveillance Center, San Diego, CA.